DOCUMENT RESUME

ED 391 381 FL 023 541

AUTHOR Chen, H. Julie
TITLE Metapragmatic Judgment on Refusals: Its Reliability and Consistency.
PUB DATE Nov 95
NOTE 43p.; Paper presented at the Annual Meeting of the American Council on the Teaching of Foreign Languages (29th, Anaheim, CA, November 1995).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Attitudes; Comparative Analysis; *English (Second Language); *Interrater Reliability; *Language Patterns; Language Research; Language Usage; *Native Speakers; *Pragmatics; Questionnaires; Written Language
IDENTIFIERS *Refusals

ABSTRACT

A study investigated 42 native English-speakers' (NSs) perceptions of the pragmatic appropriateness of refusal statements. The NSs rated the appropriateness of 24 written statements in 4 different refusal scenarios, which were collected from both native speakers and non-native speakers. Four weeks later, as a reliability check, the subjects rated the same statements again. Results indicate that what one subject considered pragmatically appropriate tended to be considered the same by other speakers of the same language. Also, the stronger the pragmatic impression, the more extreme the ratings, and the higher the level of rating consistency for a statement. In addition, subjects' pragmatic judgments tended to be consistent over time, and statements made by native speakers of English were considered pragmatically more appropriate than statements made by non-native speakers as judged by native raters. Appended materials include the questionnaire used and data charts and summaries. Contains 24 references. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 391 381

Metapragmatic Judgement on Refusals: Its Reliability and Consistency

H. Julie Chen
Indiana University

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

H. Julie Chen

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Running head: METAPRAGMATIC JUDGEMENT ON REFUSALS

FL023511

BEST COPY AVAILABLE

Abstract

Research has compared native and non-native pragmatic patterns from various perspectives, using multiple methods. No one, however, seems to have turned to the speaker of the language and ask him/her directly *how do you FEEL about the way the statement is made?* Given the belief that pragmatics is organic and intuitive, the pragmatic appropriateness of a statement is most precisely determined by native speakers intuitively and holistically rating it. In this study, 42 native speakers were asked to rate the pragmatic appropriateness of 24 written statements in four different refusal scenarios, which were collected from native and non-native speakers of English (Chen, 1991). Four weeks later, as a reliability check, the subjects rated the same statements a second time.

Results show that 1) what one considers pragmatically appropriate tended to also be considered the same by other speakers of the same language. Also, the stronger the pragmatic impression, the more extreme the ratings, and the higher the level of rating consistency for a statement; 2) subjects' pragmatic judgements tended to be consistent over time; and 3) statements made by native speakers of English were considered pragmatically more appropriate than statements by non-native speakers as judged by the native raters.

Native-Speakers' Intuition: Rating Pragmatic Appropriateness

1. INTRODUCTION

Applied linguists distinguish non-native speakers' pragmatic competence from their grammatical competence; pragmatic competence relates to the ability to function socioculturally felicitously, while grammatical competence alludes to language abilities such as syntactical and phonological competencies. It has been understood that non-native speakers' pragmatic competence does not always develop with the acquisition of their grammatical competence. Hartford and Bardovi-Harlig (1992a) indicate that being able to function linguistically (i.e. grammatically) in a foreign/second language (L2) does not necessarily mean being able to use the target language in a socioculturally appropriate manner. Because of the discrepancy between the two types of competencies, nonnative speakers, even those who are highly linguistically proficient, often do not know how to follow context-specific, sociocultural rules; they may make a statement that is perfectly grammatical but socioculturally inappropriate under a given situation or fail to comprehend what is pragmatically meant by what is literally said in cross-cultural interaction and therefore result in pragmatic inappropriateness, or *pragmatic failure* (Thomas, 1983).

Speech acts (e.g., apologies, refusals, appreciations, compliments, etc.,) since they are often socioculturally and sociolinguistically embedded, have been considered good ways to gain insight into one's pragmatic competence and have, therefore, been studied by applied linguists and L2 educators. Speech act data gathering methods that are frequently used in research include natural data observation, role play, and Discourse Completion Tasks. Since these methods have varied limitations, this study employed a different data gathering method by asking 42 undergraduate native speakers of English to holistically rate on a five-point Likert scale the pragmatic appropriateness of refusals carried out by both native and non-native speakers of English. The assumption of this study is that judgement of pragmatic appropriateness is intuitive and thus can be best projected by native speakers of the language reacting intuitively and holistically to speech act statements. The rationale is to investigate 1) whether what one considers pragmatically appropriate is also considered by other speakers of the same language to be equally pragmatically appropriate, 2) whether one's pragmatic judgement is consistent over time, and 3) whether statements made by native speakers tend to be considered pragmatically more appropriate than statements by non-native speakers as judged by native raters. Results in general indicate that one's pragmatic judgement tended to be consistent over time, that there was a high consistency of

pragmatic judgement across raters speaking the same language, and that pragmatic ratings were higher for native statements than for non-native ones.

2. BACKGROUND

2.1 Pragmatic Competence

2.1.1 Form vs. Meaning/Function

Leech (1983) divides linguistics into "grammar" and "pragmatics." The former refers to the decontextualized formal system of language, while the latter alludes to the use of language in a goal-oriented speech situation in which the speaker uses language to produce a particular effect in the mind of the hearer. Similarly, Thomas (1983, p. 92) considers linguistic competence to be made up of two types of competencies: grammatical competence (including the abstract or decontextualized knowledge of intonation, phonology, syntax, semantics, etc.) and pragmatic competence (i.e. the ability to use language effectively in order to achieve a specific purpose and to understand language in context). As the grammatical part of language is associated with the form of language whereas the pragmatic part is concerned with meaning/function, and as one form may carry several levels of meanings or serve more than one function, a mismatch between form and meaning/function results in a situation in which the "sentence meaning" differs from the "speaker's meaning" (Leech, 1983), or what Thomas calls "pragmatic failure" (1983).

2.1.2 Pragmatic Failure

Pragmatic failure is not necessarily a result of grammatical deficiency of the speaker; in Hartford and Bardovi-Harlig's study (1992a), pragmatic failure is observed in grammatically highly proficient non-native speakers of English. Rather, pragmatic failure is considered to result from the differences between languages in social rules of speaking (Hymes, 1972; Wolfson, 1983b) and from "pragmatic transfer" (Leech, 1983), when non-native speakers transfer the pragmatic norms of their community to the use of the target language. Given a universal principle of speaking across all languages, there exists a culture- and language-specific pattern of realization in each language. Take, for example, the principle of "face" (Brown and Levinson, 1978). In communication and interaction, two aspects of people's feelings are involved with "face." One is the desire of the individual "not to be imposed on," which is the "negative face," and the other, the "positive face,'" is the desire of the individual "to be liked and approved of." It is universal that all languages observe this "face" principle by saving both the speaker's and the listener's positive or negative face. The actual way in which this "face" principle is localized or realized is language specific. Pragmatic failure tends to occur when a speaker goes across languages to observe a universal principle in a non-native language which realizes the principle very differently from his or

her native language. In observing the "face" principle, for instance, one language may require a degree of directness much higher or lower than another language does. Therefore, a speaker, in preserving "face," may fail to notice the difference in the degree of directness in the two languages and thus sound too direct or indirect in the other language.

2.2 Speech Acts

Like the aforementioned principle of "face", speech acts, although they exist universally in all languages, can be highly culture- and language-specific due to the different ways they are carried out in different cultures or languages. Consequently, in investigating cross-cultural, non-native pragmatic competence, researchers often focus on speech acts and compare native and non-native speech act patterns. Schmidt and Richards (1980) maintain that "speech act theory can contribute to our understanding of second language acquisition" (p. 141). Hatch (1992) also alludes to the pedagogical implication of speech act research, "speech act theory has led to the design of the notional-functional syllabus in language teaching, setting off a major change in language teaching methodology -- away from an emphasis on the linguistic form to an emphasis on language use" (p. 136).

Much research has been done on major speech acts and various factors (e.g. familiarity, gender, and relative social status of the interlocutors) have been used to analyze or account for differences observed between native and non-native speech act utterances. Speech acts that have been studied to compare native and non-native pragmatic competence include expressions of gratitude (Eisenstein & Bodman, 1986,) apologies (Borkin & Reinhart, 1978; Cohen & Olshtain, 1981), request (Blum-Kulka, 1982), complaints (DeCapua, 1989; House & Kasper, 1981), compliments (Wolfson, 1983a, 1989b), suggestions (Bardovi-Harlig & Hartford, 1990,) and refusals (Beebe & Cummings, 1985; Takahashi & Beebe, 1987; Beebe et al., 1990; Bardovi-Harlig & Hartford, 1991; Hartford & Bardovi-Harlig, 1992b).

Among these speech acts, refusals are considered to be a face threatening act (FTA) in that either the speaker's or the listener's positive or negative face is risked when a refusal is called for or carried out. Consequently, refusals, as a sensitive, subtle, and high-risk FTA, can provide much insight into one's pragmatics since the high level of pragmatic competence required in refusals distinguishes native and non-native pragmatic competence.

2.3 Data Gathering Methods

Pragmatic data can be gathered in many different methods, each of which has different limitations. In general, data can be collected by observation of real-life events, role-plays, written questionnaires (or Discourse Completion Tasks,) or metapragmatic judgement tasks.

2.3.1 Authentic Data

Since pragmatics is highly context-driven, it is best illustrated when a speech act actually occurs and is immediately recorded as natural, authentic data, which document specific utterances (including pitches, tones, paces, pauses, etc.) with detailed description of the situation, the event, and the non-verbal reaction, status, gender, age, and relationships of the interlocutors. The greatest strength of authentic data is that they have high internal validity as they are events that actually happened and are described in detail objectively. They also provide a rich context so that a researcher or reader, when judging pragmatic appropriateness of the interlocutors, can reconstruct the entire speech act event based on the natural data and the likelihood of his or her misinterpreting the situation is thus greatly reduced.

However, natural, authentic data are very hard to come by as speech act events occur unpredictably and hence they are rarely recorded. In addition, it is highly unlikely that, in real-life situations, a given speech act occurs with the same event, in the same context, and/or with interlocutors' of the same relationships. Because the variables in real-life speech act events not only are very complicated but also can hardly be held constant in order to make cross-event comparisons, speech act events can only be studied and analyzed as individual cases; generating an abstract principle out of limited, individual, different situations can be difficult due to the frequent lack of a common base of comparison and therefore authentic data tend not to have high external validity.

2.3.2 Role Play

Since authentic data are hard to come by, role play is sometimes used to collect naturalistic data by having subjects act out plausible pragmatic interactions under a given situation. Its advantage is that the pragmatic interactions observed in role play are contextualized; since role play often takes several turns in the discourse, it can provide much insight into the meaning negotiation process.

However, one primary drawback of role play as a data collecting method is that subjects tend to be obliged to produce the item the investigator is interested in studying (Larsen-Freeman and Long, 1991, p. 27) and thus result in Hawthorne effect. In addition, in doing role play, subjects understand that the situation is not a real-life one and there is no risk of threatening either the speaker's or the listener's positive or negative face. Consequently, they may, to a certain extent, exaggerate the pragmatic interaction just to create a role-play, dramatic effect and produce a pragmatic interaction substantially different from that in a real-life situation.

2.3.3 Discourse Completion Tasks

In a Discourse Completion Task (DCT), scenarios that call for specific speech acts are presented to subjects in written form; subjects respond in writing what they think they would

actually say under the situations described in the scenarios when they are to carry out the speech act. Devices such as semantic taxonomy (Beebe et al., 1990) are then used to analyze DCT data by categorizing responses into semantic units to discover and compare native and non-native speech act patterns.

DCT enables the collection of a large amount of data so that bias can be reduced. In addition, because all subjects are presented with the same scenarios and respond in written form, data analysis tends to be more consistent and reliable. Beebe and Cummings (1985) have demonstrated that, for refusals, subjects' intuitions about what they would say correspond closely to what other subjects actually did say in the same situation. However, because real-life verbal interactions involve much more elaboration, especially in face-threatening situations, using DCT as a tool to investigate speech acts has been somewhat controversial. Wolfson, Marmor, and Jones (1989) maintain that it is not reasonable to always assume that written responses are representative of spoken ones and that short, decontextualized written segments are not comparable to authentic, longer routines (p. 182). On the other hand, Beebe and Cummings (1985) claim that written responses are stream-lined, containing fillers for slots that the subject feels must necessarily be filled to serve a particular function (p. 201) and therefore adequately capturing the essence of their oral counterparts. In spite of the controversy, DCT is still the most frequently used method in the field, as authentic data of a specific speech act is hard to collect and compare and role play does not result in comparable data as DCT does.

2.3.4 Metapragmatic Judgement Task

Few studies have employed metapragmatic judgement tasks, possibly because of its difference from other data gathering methods that there has to be existing speech act data to enable the judgement task. In other words, a metapragmatic judgement task does not generate speech act data but assesses them. It can be used as an interpretative tool with the aforementioned data gathering methods. In this study, a metapragmatic judgement task is employed. The assumptions are that 1) pragmatic impression is intuition-driven and the pragmatic appropriateness of a given statement should be intuitively and holistically judged by the speakers of that language, and 2) pragmatics is context- and discourse-based and a speech act statement is often more than the total of the semantic units that it is composed of; breaking down a speech act statement for analysis may alter or distort its pragmatic nature.

The metapragmatic judgement task is employed in this study because it lends itself to the three research questions of this study:

1) Whether what one considers pragmatically appropriate is also considered by other

speakers of the same language to be equally pragmatically appropriate,

2) Whether one's pragmatic judgement is consistent over time, and
3) Whether statements made by native speakers tend to be considered pragmatically more appropriate than statements by non-native speakers as judged by native raters.

Apparently, the greatest strength of a metapragmatic judgement task seems to be that it appeals to the subjects' (i.e. raters') intuition, which is the foundation of pragmatics and thus has high validity. That pragmatic intuition is holistic and subjective also makes a metapragmatic judgement task a reliable measurement instrument for the construct. In addition, it allows the comparison of intra- and inter-rater pragmatic judgement consistency and thus enables the investigation of whether pragmatic appropriateness of a statement as judged by a rater is consistent over time and also conforms to that by other raters. High consistency of intra- and inter-rater pragmatic judgement can in turn serve as a reliable indicator for the level of pragmatic appropriateness of individual speech act statements and thus allows the comparison of native and non-native pragmatic appropriateness.

3. METHODS

3.1 Subjects

Forty-two (20 male and 22 female) native English-speaking college undergraduate students participated in this study. These subjects are of similar age rank and share common cultural experience. Because the study was conducted in two different times, the subjects, according to their schedule, fell into two groups -- Group A (13 subjects) and Group B (39 subjects).

3.2 Instrument -- Questionnaires

Three questionnaires (i.e. I, II, & III) were used in this study. The questionnaires were based on a study conducted earlier (Chen, 1991), in which four scenarios that call for the speech act of refusals were given to 26 native and non-native speakers of English. Respondents were graduate students of different majors in Indiana University. The non-native subjects included 7 Chinese, 5 Japanese, 4 Koreans, and 5 Europeans, and there were 5 native subjects. These subjects responded to a DCT in which there were eight scenarios representing four types of refusal-eliciting stimuli (i.e. requests, invitations, offers, and suggestions (adopted from Beebe et al., 1990).) Cultural bias was minimized by carefully selecting and wording the scenarios. These scenarios and the subjects' responses in the DCT were randomly selected to compose Questionnaire I (see Appendix A) of this study.

In Questionnaire I, there were four scenarios with six responses for each. Of the six

responses, two were made by native speakers and four by non-native speakers of English in the 1991 study. The nationalities and the language background of the respondents providing the statements were not available on the questionnaires. Neither was there any indicator as to whether the responses were made by native or non-native speakers. A five-point Likert scale was given to the subjects in this study to rate the appropriateness of each speech act statement (i.e. response) in the scenarios. The rating ranged from "very inappropriate" as "1" to "very appropriate" as "5" on the scale.

The subjects were instructed to take an "outsider" perspective (as opposed to an interlocutor's perspective) and rate the appropriateness of each statement under the situations described in the scenarios. Specifically, the subjects were told that "*Imagine that you happen to eavesdrop or witness a situation as any of those scenarios on the questionnaire, and someone responds with something as any of those responses under the scenarios. As an outsider, what would you think the appropriateness of each of the statement is under the situations?*" The subjects were given an outsider's perspective in order for them to give subjective ratings without being involved in the situations so as to avoid any bias effect. If the subjects were involved in the scenarios as the interlocutors, they might translate the scale into "how likely or often would *I* make such a statement under the situation." In this case, feeling that they might be judged personally by the ratings they gave, they would want to project a best possible self-image by not giving their authentic opinions.

Four weeks later, the subjects were given a post-test with the same instruction. Some subjects recognized immediately that the scenarios and statements in the posttest were the same as those in the pretest but were still instructed to read through the scenarios and truthfully rate the statements according to their opinions.

In the posttest, subjects in Group A were given Questionnaire II and subjects in Group B Questionnaire III. (See Appendix B for the illustration of the research design.) Questionnaire II and III are composed of the same scenarios and speech act statements as those in Questionnaire I but with different scenario or statement orders. In Questionnaire II, the internal order, or order of speech act statements of Scenario II and IV, was randomized within the two scenarios. In Questionnaire III, the external order, or order of the four scenarios, was randomized. The purpose of randomizing the statement- and scenario-order in Questionnaire II and III was to investigate whether there was an order effect (i.e. the order in which the subjects respond to the statement or scenario effects the ratings.)

3.3 Data Analysis and Discussion

All of the subjects' pretest and posttest data were entered and pre-post differences were

calculated (see Appendix C). The randomized orders in Questionnaire II and III were matched with the scenario- and statement-orders in Questionnaire I. In the following discussion, the scenario and statement numbers refer to those in Questionnaire I. For instance, IV-1 refers to Scenario IV, Statement (a) in Questionnaire I.

There were five foci in analyzing the pretest and posttest data:

1) Frequency distribution of each statement under each scenario
2) Statement means and standard deviations
3) Consistency of pre-and-post ratings for individual statements across subjects
4) Consistency of pre-and-post judgement by individual subjects across statements
5) Order effect

3.3.1 Frequency Distribution

Frequency distribution was calculated for each statement (see Appendix D.) Two distribution patterns emerged for the 24 statements of the four scenarios:

a. *Clustered Distribution*:

The frequency distribution tended to cluster or peak towards either the higher or the lower end of the Likert scale for the majority of the statements. These statements include I-1, I-2, I-3*, I-4, I-6*, II-1*, II-2, II-4*, II-5, II-6, III-1, III-2, III-3*, III-4, III-5*, III-6, IV-1, IV-2*, IV-3, IV-4*, & IV-6 (* denotes statements made by native speakers of English in the 1991 study).

Depending on which end of the scale they cluster towards, some of these statements can be furthered categorized into two types: high-appropriateness cluster and low-appropriateness cluster. All of the statements made by native speakers of English show a high-appropriateness cluster, whereas statements with a low-appropriateness cluster were made by non-native speakers of English (including I-2, I-4, II-2, II-5, II-6, III-4, and III-6.)

b. *Scattered Distribution*:

There are a few statements that either do not have as sharp a peak or lack a clear tendency on the distribution. These statements are I-5, II-3, and IV-5.

3.3.2 Means and Standard Deviations

Appendix E shows the mean and the standard deviation of each statement. As aforementioned, within each scenario, almost all the statements made by native speakers have

higher means of appropriateness in comparison with statements made by non-native speakers. In examining all the native statements, it seems that what makes these statements highly appropriate are that 1) Native statements attribute the reasons of their refusals to some ***external, pre-existing, impersonal,*** or ***uncontrollable*** situations, such as "I am going to a concert and we bought the ticket a long time ago" (II-4*) in turning down a dinner invitation, and "I need the car/notes myself" (I-3*, IV-2*, & IV-4*) in turning down a request for borrowing the car/notes. By externalizing or impersonalizing the reasons of refusals, the speaker avoids threatening his or her positive face (i.e. one's desire of being liked) and/or the listener's negative face (i.e. one's desire of not being imposed upon); 2) Native statements ***elaborate*** the reasons of refusals. Perhaps attempting to reduce the intensity of the situation as caused by the refusals and to support the reasons, native refusals in general are lengthier and better elaborated than non-native refusals; 3) Native statements offer (well-intended) ***alternatives*** or ***suggestions*** and thus open up other doors for the listener, instead of abruptly stating "no". Some alternatives or suggestions are "(I can't lend you all my notes, but) is there one particular class that you need some notes on?" (I-6*) "Maybe someone in your study group can loan you the notes," (I-3*) "...But, could we go to lunch Tuesday together?" (II-1*) and "Perhaps we can get together some other time" (II-4*); and 4) Native statements down grade the refusals, by ***minimizing the undesirable outcome*** that the listener caused and offers to make up for. For instance, "Don't worry. *Really*..." "...It's *nothing* at all. I'll get it *in a minute*." (III-3*) as refusals to the listener's offer to help cleaning up the spilt coffee and thus make the listener feel "off the hook."

In contrast, standard deviation is low for several non-native statements that also have low appropriateness average, suggesting that these non-native statements are very consistently considered by most raters to be very inappropriate. Specifically, these non-native statements are "I think you should think about what you are doing" (I-2) as a refusal to lend out notes, "Is your husband/wife going to be there? If so, no. Thanks." (II-5) and "Why don't you get a divorce?" (II-6) as refusals to a friend's invitation to dinner, and "Let me do it. You'll only make matters worse" (III-4) as a refusal to the offer of helping to clean up. What seems to be common about these highly inappropriate statements is that imbedded in the refusals are additional messages that criticize the listener personally, with an overtone suggesting the reasons of refusals being something negative about the listener, similar to "you should be *ashamed* that you not only didn't come to class but also want someone else's notes that you didn't work for," "You are married to a *wrong* person," and "You are so *clumsy* that I can't let you help me clean up the mess you made." It is difficult enough to

preserve the speaker's positive face when carrying out a refusal; the speakers in the above situations take even one step further to threaten the listeners' negative face by making personal criticisms. Failure to maintain the ***listener's negative face*** in the native way is one cause of the pragmatic inappropriateness in these non-native statements. Native strategies to save the listener's negative face include externalizing / impersonalizing (as opposed to internalizing / personalizing) the situation and minimizing the undesirable outcome caused by the listener (as opposed to emphasizing the possible negative consequence). The best example to illustrate this point is "Let me do it. *You*'ll only make matters *worse*." in the spilt coffee scenario (III-4, mean=1.64.) The speaker of this statement personalizes the refusal (cf. "This kind of things happen. Things break," (III-2, mean=4.01) which externalizes and impersonalizes the situation.) In addition, the speaker of III-4 also emphasizes the negative consequence (c.f. "It's nothing at all..." (III-3*, mean=4.01)) and appears to have failed to save the listener's negative face.

Also, a closer examination of the aforementioned four highly inappropriate non-native statements (i.e. I-2, II-5, II-6, and III-4) suggests that not only do they fail to employ (in the native way) the appropriate refusal strategies and to preserve the listener's negative face, some of them further aggravate the inappropriateness by making personal criticisms under an inferior ***status of favor***. The violation of status of favor is another cause of pragmatic failure for some of these highly inappropriate statements. These four statements illustrate two situations: refusals to someone's request (e.g., borrowing notes/car) and refusals to someone's offer (e.g., to dinner or to help cleaning up). These two situations differ in that, in the first situation, the person making refusals is a possible favor-granter and thus is in an advantaged or favor-superior position, while in the second situation, the person making the refusal is a possible favor-receiver and thus is in an disadvantaged or favor-inferior position. It is already presumptuous for someone in an inferior, favor-receiving position to refuse, and thus threaten the positive face of the superior favor-granter; the inappropriateness is even further aggravated by the favor-receiver turning down the well-meaning favor-granter with personal criticisms or attacks (e.g., suggesting a divorce, which differs very much from the well-intended alternative/suggestion strategy observed in the appropriate, native statements). The observation that the status of favor is a factor of pragmatic appropriateness is supported by the fact that the refusals made by the inferior-status-of-favor threatening the negative face of the superior-status-of-favor (i.e. II-5, II-6 and III-4; refusals with personal criticisms to dinner invitation and to offer cleaning up) have the lowest appropriateness means (1.61, 1.04 and 1.64 respectively) of all statements. In comparison with these three refusals by inferior

status of favor, the refusals by superior status of favor (I-2), though highly inappropriate, still has a significantly higher mean of appropriateness (2.71), because, although the listener's negative face is threatened by the speaker's personal criticism, the speaker has more of the right to do so as he/she is in a superior status of favor.

A comparison of III-5 and III-6 (both are refusals to offer help to clean up spilt coffee) renders insight into another possible cause of pragmatic inappropriateness. The two statements are similar in that they both externalize the reasons of refusals ("I have some handy cleaning stuff..." in III-5 and "I'm going to buy the cleaning kit later" in III-6.) However, one (III-5) results in a high appropriateness mean (4.17) and the other (III-6) a much lower mean (2.62). The difference seems to lie in the degree of semantic clarity of the reason of refusals. III-5 expresses clearly that the speaker has a handy stuff for the problem and thus refuses the listener's offer to help, while III-6 does not make as clear a refusal (i.e., buying the cleaning kit later does not readily translate into an acceptance or a rejection of the listener's offer to help.) This seems to suggest that, although not as detrimental, ***semantic ambiguity*** can be a cause of pragmatic inappropriateness. Another example of semantic ambiguity causing pragmatic inappropriateness is IV-5 (mean=2.63), "I *might* go somewhere. I'll give you may answer *later*" as a refusal to lending out the car (cf. semantic clarity in "I *need* it *this week*" in IV-2* and IV-4*, or even "I *can't* lend you my car" in IV-6, all of which have a higher mean than that of the semantically ambiguous statement, IV-5).

3.3.3 Statement Rating Consistency

Level of consistency for the rating of each statement across raters is also calculated. The Statement Rating Consistency table in Appendix F shows that, on average, more than half (i.e. 55.3%) of the posttest statement ratings are identical with their corresponding pretest ratings (i.e. ±0, or differ by 0 point on the five-point scale), a considerable percentage (i.e. 35.5%) of the posttest statement ratings differ from their corresponding pretest ratings by only 1 point (i.e. ±1) on the five-point scale, and only 8% of the posttest ratings differ by 2 points (i.e. ±2) or more on the five-point scale from their corresponding pretest ratings. In other words, the statements, on average, have a very high pretest-posttest rating consistency across subjects, which, in turn, suggests a very high inter-rater reliability.

Examination of individual statements on the table shows that several of them have even higher pretest-posttest consistency (II-6 and III-3*, with 93% and 74% respectively for ±0 difference). Also, at the ±0 level, there are three statements that have significantly lower-than-average percentages. These are I-4, I-5, and II-3, with 33%, 40% and 40% respectively at the ±0 level. However, these three statements "catch up" on their level of

pretest-posttest rating consistency with significantly higher-than-average percentages (48%, 43%, and 40% respectively) at the ±1 level, suggesting that, although these three statements have relatively lower percentages of identical pre-post ratings, they are not "far off" as over 80% of the ratings in each of these three statements differ by 1 point or less in the posttest.

On the other hand, investigating the pragmatic nature of those statements that have significantly higher or lower percentages of pre-post rating consistency may provide us with some insight into what influences pre-post rating consistency. It seems the three statements (I-4, I-5, and II-3) that have lower-than-average pre-post rating consistency at the ±0 level have in common that they are semantically vague and can be pragmatically translated in many ways and thus result in lower pre-post rating consistency. The statement that "If you had been in class, they (the notes) would make sense to you" (I-5) does not really tell the listener whether he/she can borrow the notes. The lack of semantic clarity in this statement grants the rater greater latitude in interpreting its pragmatic message and this may be the reason why this statement also has a scattered distribution and a high standard deviation (1.11). The statement that "You should've taken notes by yourself!" (I-4) is similar to I-5 in the same way. This observation that semantic clarity influences the consistency of pragmatic rating is in line with the aforementioned distinction Leach (1983) draws between "sentence meaning" and "speaker's meaning;" the incongruity of these two meanings, caused by the lack of clarity in "sentence meaning" in the above examples, results in a higher latitude for pragmatic interpretation and thus a lower consistency in pre-post rating. The mismatch between the two types of meanings also conforms with what Thomas (1983) calls "pragmatic failure" and explains why these statements have somewhat lower means (2.64 for I-4 and 2.94 for I-5) of pragmatic appropriateness.

II-3 is the third statement that has a lower-than-average pre-post consistency percentage at the ±0 level. This statement seems to project a different reason for lower pre-post consistency. The reason, instead of semantic ambiguity, is an opposite one in that it states the ***truth*** (that the speaker does not like the spouse of the listener) in an ***honest*** way, assuming that the listener already knows the reason ("*You know* that I don't get along with your husband/wife") and straightforwardly suggests the possible ***negative consequence*** of the event ("It'll be awkward when we are together") as the reason for the refusal. Apparently, there is a high congruence between the sentence meaning and the speaker's meaning in this statement. What causes it to have a scattered distribution, a high standard deviation, and a relatively low pre-post rating consistency may be the different pragmatic strategies (truth, honesty, expressing negative consequence) that can be highly individual-

case-based and that people in general do not have a strong preference for or prejudice against. Thus, whether these strategies are positive or negative by themselves and whether the positive virtue of truthfulness and honesty in the statement outweighs the negative impression as caused by lack of tactfulness and well-roundedness in the way the truth is told are very much left to a rater's personal judgement. Given the many complicated, and even contradictory qualities of this statement, it is likely that the rater may have a mixed pragmatic feeling about the statement and therefore gives it a considerably different rating of pragmatic appropriateness in the posttest, resulting in the statement's lower pre-post rating consistency at the ±0 level.

Although not a statement of low pre-post rating consistency at the ±0 level, statement IV-3 forms an interesting comparison with II-3 and hence helps to more closely examine the interaction of the many qualities in II-3. IV-3 is similar to II-3 in that it also expresses a negative consequence as the reason for refusing to lend out the car ("If something happens, it's a mess with the insurance.") Whereas both II-3 and IV-3 have a mean indicating positive pragmatics (3.1 for II-3 and 3.75 for IV-3), IV-3 is undoubtedly pragmatically more appropriate than II-3. This difference may be attributed to the fact that IV-3 expresses the negative consequence (that the insurance can be messed up) as the reason for refusal *without* alluding to the truth (that the listener is a bad driver,) which would threaten the listener's negative face. On the other hand, in II-3, although telling the truth (that the speaker does not get along with the friend's spouse) is a virtue and can be considered a positive strategy, it nonetheless violates the rule that the listener's negative face should be saved.

In addition, all these three statements (I-4, I-5, and II-3) that have a lower-than-average pre-post rating consistency at the ±0 level also have means very close to the middle of the five-point scale (i.e. "3".) The middle (or "mild") ratings indicate that raters either did not have a strong opinion or have varied opinions about these statements.

As for the statements (II-6 and III-3) that have extremely high pre-post rating consistency at the ±0 level, although one pragmatically negative (mean=1.03 for II-6) and the other positive (mean=4.01 for III-3), they have in common the clear tendency of high (in)appropriateness; their means fall at the ends of the five-point scale and are either highly pragmatically appropriate or extremely pragmatically inappropriate. This tendency is also illustrated by their frequency distributions having sharp peeks at the higher/lower ends.

This observation that highly (in)appropriate statements tend to have strong pre-post rating consistency (at the ±0 level), in conjunction with the previous discussion about semantically vague or pragmatically contradictory/complicated statements tending to have a

middle/mild pragmatic rating and thus a lower pre-post rating consistency (at the ±0 level), indicates that there does exist a correlation between the level of pre-post rating consistency (at the ±0 level) and the "strength" of the positive or negative pragmatic feelings a statement renders the raters. A rater, when giving a very high or low pragmatic rating to a statement, tends to have a strong opinion about the pragmatics of the statement, and, since strong opinions tend to stay stable, the rating the rater gives to that statement in the posttest tends to have high consistency with that in the pretest. This correlation is quite obvious as illustrated in Figure 1 below:

Insert Figure 1 about here

3.3.4 Subject Judgement Consistency

Appendix G shows the table of each of the 42 subjects' pretest-posttest judgement consistency across statements as projected by their pragmatic ratings. On average, 52.6% of the subjects' pragmatic judgement in the posttest matches identically (i.e. ±0) with their pragmatic judgement in the pretest. A considerable percentage (34.9%) of the subjects' posttest pragmatic judgement differs from their pretest pragmatic judgement by only 1 point (i.e. ±1) on the five point scale. Less then 13% of the subjects' posttest pragmatic judgements differ from their pretest pragmatic judgements by 2 points (i.e. ±2) or more on the five-point scale, and 2.5% differ by of 3 points (i.e. ±3) or more. There are a couple of subjects that have relatively low pre-post pragmatic judgement consistency at the ±0 level (33% for Subject B8 and 29% for Subject B21). However, their judgement consistency at the ±1 level also indicates that they "catch up" on their level of judgement consistency with a significantly higher-than-average percentage (50% for Subject B8 and 58% for Subject B21), so that, combining both the ±0 and the ±1 levels, they still manage to maintain a percentage higher than 80% of ±0 or ±1 pre-post judgement consistency, and this percentage is by no means significantly different from the average (i.e. 87.5% as the total of the ±0 and ±1 levels).

The subjects' high consistency of pretest-posttest pragmatic judgement indicates a high intra-rater reliability.

3.3.5 Order Effect

It might be possible that the first scenario/statement that the subjects responded to

"anchored" the ratings for the ensuing scenarios/statements, which would be compared to and influenced by how high or low the rating of the first scenario/statement is. It might also be possible that, since the scenarios and statements in the posttest are identical with those in the pretest, the subjects might have recalled during the posttest the ratings they gave in the pretest not (only) because of the actual pragmatic qualities of the statements but the order in which the scenarios/statements were presented to them in the pretest. Consequently, order effect is examined by comparing the three questionnaires (which only differ in their scenario or statement order). The Order Effect table in Appendix H shows that, in comparing Questionnaire I with Questionnaire II (which differ in the internal/statement order within Scenario II and IV,) Scenarios I and III, whose statement orders remain the same in the two questionnaires, do not yield a total of rating difference significantly different from those of Scenario II and IV, whose statement orders are randomized in Questionnaire II. Because of the lack of significant difference between the randomized and the unchanged statement orders, it seems that there is no order effect in the questionnaires and that the ratings the subjects gave to the statements are independent from the order in which the statements were presented to the subjects.

In addition, because of the absence of order effect, as observed among subjects in Group A (who had Questionnaire II in posttest), it is assumed that there is also no order effect in Questionnaire III, which were given to subjects in Group B in the posttest with the scenario order randomized.

4. RESULTS

The above observations from the data help answer the three research questions projected earlier. The questions are:

1) ***Is what one considers pragmatically appropriate also considered by other speakers of the same language to be equally pragmatically appropriate?***

The answer seems to be yes. The strongest evidence is that there is a very high statement rating consistency; in addition, the majority of the statements show clustered distributions, suggesting that the raters tended to have similar, if not identical, opinions regarding the pragmatic appropriateness of the statements.

The data analysis also suggests what the raters consider to be pragmatically appropriate and what inappropriate. The pragmatic strategies that are found among statements rated pragmatically appropriate include:

- *Externalizing or impersonalizing the situation of refusals,*

- *Elaborating the reasons of refusals,*
- *Offering alternatives or suggestions, and*
- *Minimizing the undesirable outcome the listener caused.*

The pragmatic strategies that are found among statements rated pragmatically inappropriate include:

- *Failure to preserve the listener's negative face,*
- *Making personal criticisms, especially under an inferior status of favor,* and
- *Semantic ambiguity.*

In addition, although the statement rating consistency is very high, the raters seemed to agree more strongly on some statements than on others. The statements that the raters showed strong rating agreement have the following characteristics:

- *They have a clear tendency to have extreme scores, either highly appropriate or highly inappropriate,* and
- *Their frequency distributions tend to cluster and have sharp peeks.*

The statements that the raters showed less strong rating agreement have the following characteristics:

- *Their scores tend to fall in the middle of the scale,*
- *They are of complicated or contradictory pragmatic qualities,* and
- *They may be semantically vague of pragmatically mild and thus they do not cast a strongly positive or negative impression on the raters.*

The rating consistency in combination with the means of the statements' pragmatic appropriateness also suggests a correlation between the strength of the positive or negative impression the statements render the raters and the consistency of the rating the statements have; the stronger the pragmatic impression, the more extreme the ratings, and the higher the level of rating consistency.

2) ***Is one's pragmatic judgement consistent over time?***

The answer for this question also seems to be yes. As discussed, the subjects' pretest-posttest judgement consistency is very high, implying that the raters tended to hold a stable pragmatic opinion over time for a given statement.

3) ***Do statements made by native speakers tend to be considered pragmatically more appropriate than statements by non-native speakers as judged by native raters?***

Once again, the answer to this question seems to be yes, which is supported by 1) common strategies or qualities are found among the native statements but among only very few non-native statements, supporting that people speaking the same language share common ways of expressing refusals, and 2) the raters in this study very consistently rated these native statements as highly pragmatically appropriate, indicating that people speaking the same language tend to agree on and approve of the ways in which other speakers of the same language commonly express refusals.

5. LIMITATIONS

Like most research on pragmatics, this study's limitations are due to its sample, the data collection method, and data analysis and interpretation.

5.1 Subjects

Since the 42 subjects participated in this study were college undergraduates with the same age range, homogeneous linguistic experience, and similar cultural background, they may very well under-represent the entire language population. However, it is difficult to determine who the representatives of a language population should be. Do they include people who have rich cross-cultural, multi-lingual experiences or even have lived in another culture and speak another language? Culturally and linguistically experienced people may be disqualified for the purpose of investigating the sociocultural pragmatic pattern of the language, as they might already be "contaminated" by the culture or language they have experience in. If the ideal subjects should be culturally and linguistically homogeneous, the subjects in this study then is quite representative. On the other hand, even among the culturally and linguistically homogeneous population, there are different variables that can influence one's pragmatic opinion. These variables may be one's age, up-bringing, gender, personality, sociocultural exposure, etc. The subjects of this study, in this sense, may fail to include many of the different variables.

5.2 Data Collection

Since the questionnaires this study used were based on a study conducted earlier (Chen, 1991), in which DCT was the format for collecting written responses, the aforementioned potential drawbacks of DCT apply to the statements in the questionnaires of this study. However, in spite of this limit, the statements in the DCT were transformed into the questionnaires without threatening their validity or distorting the data, since both the subjects in the 1991 study who produced the statements and the subjects in this study who rated the statements responded to/in the same written form.

In addition, although four weeks is a reasonable amount of pretest-posttest interval, it is hard to judge whether that is a time span long enough for the subjects to forget the numbers they associated the individual statements with in the pretest in order to react to the statements purely intuitively and pragmatically in the posttest.

5.3 Analysis and Interpretation

Because the interpretation of the data is based on the limited number of the statements (total 24 statements for four scenarios,) the support and generalization of the findings are limited. However, there is also a limit as to how many statements the raters may have patience and attention for. Another issue is the insufficient triangulation for the validity of the findings and the interpretation. Although some triangulation was found among the statements regarding what seemed to be the characteristics of appropriate pragmatics and what inappropriate, another person investigating the same data might have a different view, suggesting the next phase of the study to be conducting interviews to gather opinions about the pragmatic appropriateness of the statements or having other investigators interpret the data.

Yet another issue concerns whether the refusals in the statements are true reasons or false excuses. Prior to this study, the questionnaires were tested and given to several people to react to. One of the people gave double ratings for several statements and indicated that one rating was for "if true" and the other "if false," with the two ratings significantly different from each other. It seemed that, at least to this person, whether a statement was actually true or false (i.e. whether the speaker of the statement was telling the truth or lying, in whatever pragmatic way) was a factor of its pragmatic appropriateness. This "true or false" issue is a problem not as much of the questionnaires or statements but of the construct of pragmatics. Even in real-life situations, rarely can we find out whether one, in saying no, is telling a true reason or a false excuse. Consequently, this "true or false" issue calls for a specific definition of pragmatics and a clear distinction between pragmatics and intention. As aforementioned, a meaning can be twofold, sentence meaning (i.e. semantic meaning) and speaker's meaning (i.e. pragmatic meaning) (Leech, 1983). When the two meanings of an utterance mismatch *without the speaker's knowing it* (i.e. the speaker means one thing and does not know he/she is heard to have said another), *pragmatic* failure (Thomas, 1983) occurs; however, when the speaker means one thing but *intentionally* says another, the issue is not of pragmatics. Going back to the "true or false" issue, the respondents' not knowing whether the speakers of the statements had a pragmatic failure or actually intended to give a false excuse assimilates real-life situations very well. As we form impressions about people

often without knowing or even thinking whether something they say is of pragmatics or intention, the subjects in this study were only instructed to pretend that they had heard the statements they rated without the investigator's alluding to or sensitizing them with the "true or false" issue.

6. CONCLUSION

The findings in this study, although they distinguish native and non-native pragmatics and thus suggest the language- and culture-specificity of pragmatics, do not by themselves generate a pattern. Given the subtlety and complexity of the construct, pragmatics needs to be investigated from different perspectives and in different formats. Individual pragmatic cases encountered in real life situations provide rich insights all by themselves, discourse completion tasks collect large amount of data, questionnaire surveys yield metapragmatic data that are consistent in form and also enable comparison and generalization, selective rater interviews further explore the raters' opinions and possibly how one's experience influences his/her pragmatic judgement, and inter-investigator triangulation validates findings from, and interpretations of, the data collected in the above various formats. Only in so doing can the nature of pragmatics be thoroughly studied and the many cross-cultural, inter-language pragmatic questions be appropriately discussed.

References

Bardovi-Harlig, K., & Hartford, B. S. 1990. 'Congruence in have and nonnative conversations: Status balance in the academic advising sessions'. *Language Learning,* 40: 467-501.

Bardovi-Harlig, K., & Hartford, B. S. 1991. 'Saying "No" in English: Native and nonnative rejections'. In L. Bouton and Y. Kachru (Eds.), *Pragmatics and Language Learning* (Vol. 2, pp. 41-57). Urbana, University of Illinois.

Beebe, L. M., Takahashi, T., & Uliss-Weltz, R. 1990. 'Pragmatic transfer in ESL refusals'. In R. Scarcella, E. Anderson, & S. D. Krashen (Eds.), *On the development of communicative competence in a second language* (pp. 55-73). Cambridge, MA: Newbury House.

Beebe, L. M., & Cummings, M. C. 1985, April. *Speech Act Performance: A Function of the Data Collection Procedure?* Paper presented in the Sixth Annual TESOL and Sociolinguistics Colloquium at the International TESOL Convention, New York.

Blum-Kulka, S. 1982. 'Learning how to say what you mean in a second language: A study of the speech act performance of learners of Hebrew as a second language'. *Applied Linguistics,* 3: 29-59.

Borkin, A., & Reinhart, S. 1978. Excuse me and I'm sorry. *TESOL Quarterly,* 12: 57-70.

Brown, P., & Levinson, S. 1978. Universals in language usage: Politeness phenomena. In E. N. Goody (Ed.), *Questions and politeness* (pp. 54-310). Cambridge: Cambridge University Press.

Chen, H. J. 1991. *Native and Nonnative Refusals.* Unpublished manuscript..

Cohen, A. D., & Olshtain, E. 1981. Developing a measure of sociolinguistic competence: the case of apology. *Language Learning,* 31: 113-134.

DeCapua, A. 1989. *An Analysis of Pragmatic Transfer in the Speech Act of Complaints as Produced by Native Speakers of German in English.* Unpublished doctoral dissertation, Teachers College, Columbia University.

Eisenstein, M., & Bodman, J. W. 1986. "I very appreciate": Expressions of gratitude by native and non-native speakers of American English. *Applied Linguistics,* 5: 167-185.

Hartford, B. S., & Bardovi-Harlig, K. 1992a. Closing the conversation: Evidence from the academic advising session. *Discourse Processes,* 15: 93-116.

Hartford, B. S., & Bardovi-Harlig, K. 1992b. Experimental and observational data in the study of interlanguage pragmatics. *Pragmatics and Language Learning, Monograph Series,* 3: 33-52.

Hatch, E. 1992. *Discourse and Language Education.* Cambridge University Press: New York.

House, J., & Kasper, G. 1981. Politeness markers in English and German. In F. Coulmas (Ed.), *Conversational Routine. Explorations in Standardized Situations and Pre-Patterned Speech: Vol. 2* (pp. 157-186). Mouton: The Hague.

Hymes, D. 1972. Models of the interaction of language and social setting. In J. J. Gumperz & D. Hymes (Eds.), *Directions in sociolinguistics: The ethnography of communication* (pp. 35-71). New York: Holt, Rinehart and Winston.

Leech, G. 1983. *Principles of pragmatics.* London: Longman.

Schmidt, R., & Richards, J. C. 1980. Speech acts and second language learning. *Applied Linguistics,* 1: 129-155.

Takahashi, T., & Beebe, T. 1987. The development of pragmatic competence by Japanese learners of English. *JALT Journal,* 8: 131-155.

Thomas, J. 1983. Cross-Cultural Pragmatic Failure. *Applied Linguistic,* 4/2: 91-112.

Wolfson, N. 1983a. An empirically based analysis of complimenting in American English. In N. Wolfson & E. Judd (Eds.), *Sociolinguistics and Language Acquisition* (pp. 82-95). Rowley, MA: Newbury House.

Wolfson, N. 1983b. Rules of speaking. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 61-87). London: Longman.

Wolfson, N. 1989. The social dynamics of native and nonnative variation in complimenting behavior. In M. Eisenstein (Ed.), *The Dynamic Interlanguage* (pp. 219-236). New York: Plenum.

Wolfson, N. Marmor, T., & Jones, S. 1989. Problems in the comparison of speech acts across cultures. In Blum-Kulka, S., House-Edmondson, J., Kasper, G. (Eds.), *Cross-Cultural Pragmatics: Requests and Apologies* (pp. 82-95). Rowley, MA: Newbury House.

Figure Caption

Figure 1. Correlation between Level of Rating Consistency and Degree of Pragmatic Appropriateness.

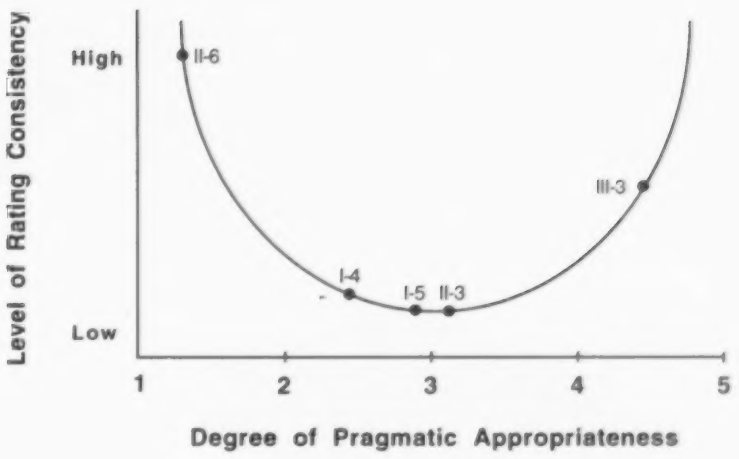


Fig. 1 Correlation between Level of Rating Consistency and Degree of Pragmatic Appropriateness

Appendix A
Questionnaire I.

Given the following four situations (I ~ IV), how appropriate do you consider each of their responses (A ~ F) to be? Give each response in each situation a rating, by circling one of the five numbers on the scale besides it.

1	2	3	4	5
very inappropriate	inappropriate	undecided	appropriate	very appropriate

I. W attends classes regularly and takes good notes. One person in W's class who doesn't show up very often asks to borrow W's notes. Since W has to compete with the rest of the class to earn a good grade, W doesn't feel like sharing the results of his/her hard work with someone who doesn't work for it. W says,

1 2 3 4 5 A) "If I lend my notes to you, it is unfair to me and others who come to class regularly."

1 2 3 4 5 B) "I think you should think about what you are doing."

1 2 3 4 5 C) "I need them to study from. Maybe someone in your study group can loan you the notes. Sorry."

1 2 3 4 5 D) "You should've taken notes by yourself!"

1 2 3 4 5 E) "If you had been in class, they (the notes) would make sense to you."

1 2 3 4 5 F) "I'm not sure my notes will help you because they relate so closely to what was said or done in class. I really would rather not have them all copied. Is there one particular class that you need some notes on?"

II. A friend invites X to dinner, but X really can't stand this friend's husband/wife. X says,

1 2 3 4 5 A) "Well, I can't that night. But, could we go to lunch Tuesday together?"

1 2 3 4 5 B) "I don't want to go out at night."

1 2 3 4 5 C) "You know I don't get along with your husband/wife. It'll be awkward when we are together."

1 2 3 4 5 D) "I'm sorry, but my husband and I are going to a concert. We bought the tickets a long time ago. Perhaps we can get together some other time."

1 2 3 4 5 E) "Is your husband/wife going to be there? If so, no. Thanks."

1 2 3 4 5 F) "Why don't you get a divorce?"

1	2	3	4	5
very inappropriate	inappropriate	undecided	appropriate	very appropriate

III. Y's friend is at Y's house for coffee. The friend accidentally spills a whole cup of coffee on Y's carpet and makes a mess of it. The friend insists on cleaning it up for Y, but Y doesn't want the friend to. Y says,

1 2 3 4 5 A) "It's my house. I don't want my guests to clean up anything."

1 2 3 4 5 B) "This kind of things happen. Things break."

1 2 3 4 5 C) "Don't worry. Really. It's nothing at all. I'll get it in a minute."

1 2 3 4 5 D) "Let me do it. You'll only make matters worse."

1 2 3 4 5 E) "No, no. Don't worry. I have some handy cleaning stuff I use for stuff like that. Sit still and keep talking. I'll get it."

1 2 3 4 5 F) "I'm going to buy the cleaning kit later."

IV. Z's roommate asks to use Z's car to go to Chicago. Knowing that the roommate is a careless and unskillful driver, Z doesn't want to lend the roommate the car. Z says,

1 2 3 4 5 A) "I'm sorry, but I don't lend my car to anyone."

1 2 3 4 5 B) "I'm sorry, but I don't feel comfortable loaning out my car for long trips. Besides, I need it this week. How would I get around?"

1 2 3 4 5 C) "If something happens, it's a mess with the insurance."

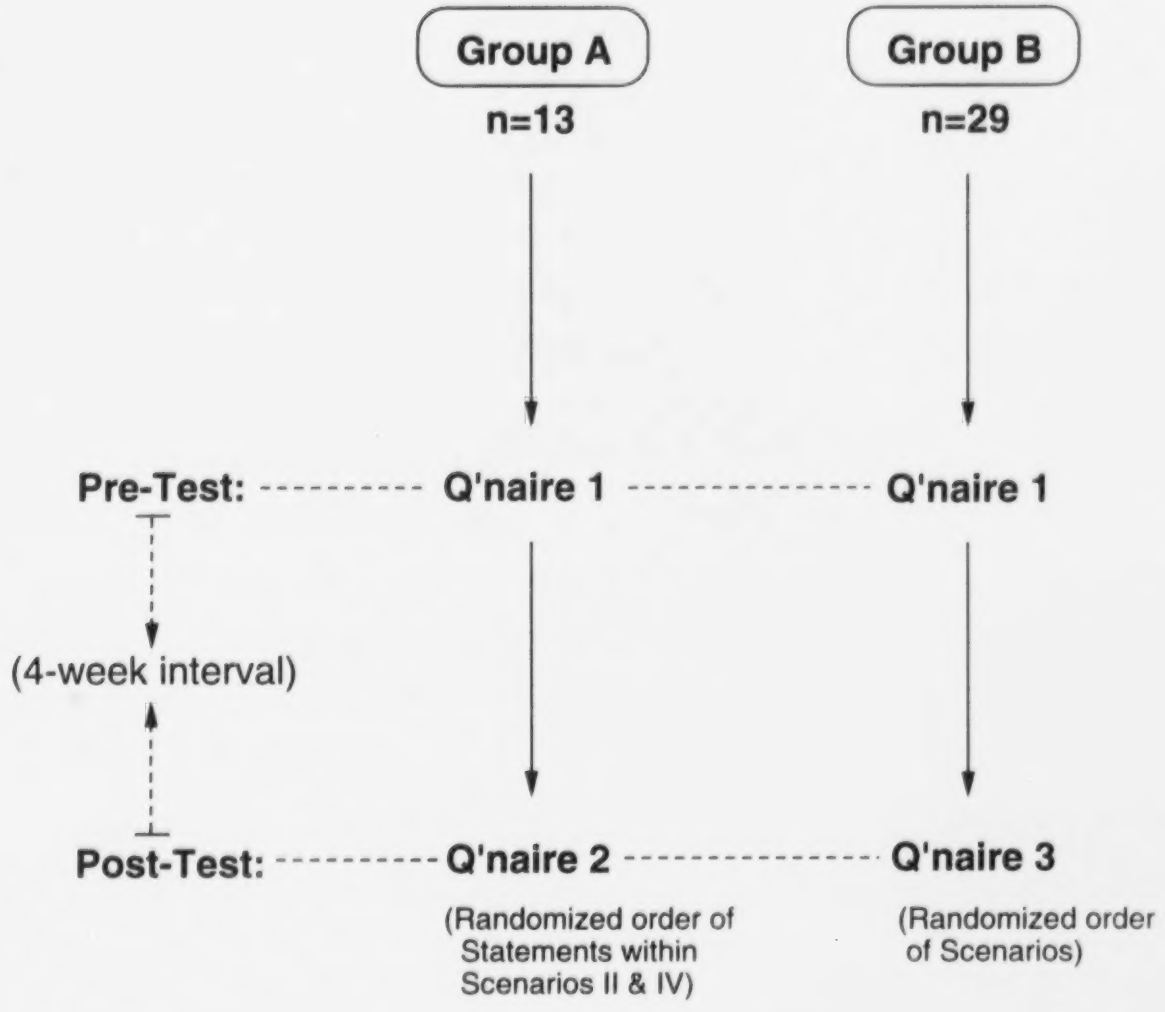
1 2 3 4 5 D) "I'm sorry, but I made a policy not to lend my car to anybody. Besides, I need it this week."

1 2 3 4 5 E) "I might go somewhere. I'll give you my answer later."

1 2 3 4 5 F) "I'm sorry. I can't lend you my car."

Appendix B

Research Design



Pretest & Posttest Scores

Subjects	I-1	I-2	I-3*	I-4	I-5	I-6*	II-1*	II-2	II-3	II-4*	II-5	II-6	III-1	III-2	III-3*	III-4	III-5*	III-6	IV-1	IV-2*	IV-3	IV-4*	IV-5	IV-6
A1-pr	4	3	4	1	3	5	4	1	1	3	1	1	1	3	4	1	5	3	3	4	4	4	2	3
A1-po	4	1	4	1	2	5	5	4	1	5	2	1	1	2	4	2	5	3	4	3	1	4	4	4
A2-pr	2	2	4	3	2	2	5	2	2	3	2	1	4	4	4	3	4	3	4	5	5	5	2	4
A2-po	3	3	4	3	3	4	5	2	3	4	1	1	4	3	3	2	4	3	4	4	5	3	3	4
A3-pr	4	3	5	3	4	5	5	2	4	5	1	1	4	4	4	2	3	2	5	5	5	5	5	5
A3-po	4	3	4	4	4	5	5	4	2	5	2	1	3	4	4	2	3	4	5	5	5	5	5	5
A4-pr	3	3	5	2	2	3	4	2	4	4	1	1	5	5	5	2	5	5	5	5	5	5	1	3
A4-po	4	3	5	4	5	3	1	4	3	4	1	1	4	4	4	2	4	4	4	5	5	4	1	4
A5-pr	5	4	5	3	5	3	3	3	1	5	1	1	4	3	4	2	4	2	3	4	2	4	3	4
A5-po	5	4	5	4	5	4	3	4	1	5	2	1	2	5	4	1	3	2	3	3	2	2	4	3
A6-pr	4	4	4	4	5	3	4	2	2	3	2	1	3	4	2	2	5	3	3	3	4	3	4	3
A6-po	4	4	4	3	4	4	5	3	2	4	2	2	4	3	2	2	4	3	3	4	3	4	3	3
A7-pr	2	5	4	4	5	3	5	2	5	3	1	1	2	5	5	1	5	2	4	5	4	3	1	4
A7-po	2	2	5	1	1	2	5	4	5	5	2	1	1	5	5	1	5	3	4	5	2	5	1	4
A8-pr	4	2	4	3	2	4	4	3	2	5	1	1	2	3	4*	1	5	3	5	4	2	4	3	5
A8-po	5	2	5	3	4	5	4	4	3	5	2	1	4	4	4	1	4	3	4	5	2	4	4	4
A9-pr	4	2	5	2	4	3	5	1	3	4	1	1	3	4	4	2	4	2	4	4	4	3	2	2
A9-po	2	4	4	3	3	4	3	3	4	4	1	1	3	4	4	2	4	2	4	4	4	3	2	3
A10-pr	4	2	4	1	3	3	5	3	3	3	1	1	5	4	5	1	5	3	4	2	4	2	2	5
A10-po	3	3	2	2	2	3	4	3	2	3	1	1	4	4	4	1	3	2	2	2	2	2	2	3
A11-pr	4	3	4	4	4	4	4	2	3	4	3	1	4	4	4	3	4	4	5	5	5	5	4	5
A11-po	3	3	4	4	4	4	4	4	3	4	2	1	4	4	4	3	4	4	4	4	4	4	4	4
A12-pr	5	4	5	2	4	5	2	4	5	5	2	1	4	4	2	1	5	3	5	5	4	4	4	4
A12-po	5	4	5	3	4	5	4	4	5	5	4	1	5	5	5	2	5	4	5	5	5	5	5	5
A13-pr	3	2	5	2	2	4	4	3	4	4	1	1	3	4	4	2	4	3	4	4	4	4	3	4
A13-po	3	2	5	2	2	4	4	3	4	4	1	1	3	4	4	1	5	4	4	5	5	4	3	4
B1-pr	4	2	4	3	4	5	4	2	4	3	2	1	2	4	4	2	5	2	5	5	3	5	3	4
B1-po	4	2	5	3	4	4	4	2	3	3	3	2	2	4	4	1	5	1	4	4	3	5	3	4
B2-pr	4	2	4	1	1	4	4	2	1	4	1	1	4	4	5	1	5	2	2	4	4	4	2	1
B2-po	4	3	4	3	2	4	4	2	1	4	1	1	4	3	5	1	5	1	3	4	3	4	2	2
B3-pr	5	2	4	3	4	2	2	1	5	2	1	1	3	1	5	1	4	2	5	4	4	3	2	5
B3-po	5	3	2	5	4	2	2	1	5	3	2	1	3	2	4	1	4	3	5	4	5	3	2	4
B4-pr	3	3	4	1	3	2	5	2	3	5	1	1	5	5	5	2	5	3	5	5	5	5	3	4
B4-po	2	1	4	3	3	1	5	2	2	5	2	2	5	5	5	2	3	2	5	5	5	5	2	5
B5-pr	4	3	4	2	3	5	4	1	5	3	2	1	4	4	3	2	4	2	3	4	5	4	3	5
B5-po	4	2	4	2	3	5	4	2	4	2	1	1	5	5	2	1	3	2	2	3	5	3	4	5
B6-pr	3	4	5	3	4	3	4	2	4	5	2	1	4	5	5	2	5	3	4	5	3	5	3	5
B6-po	4	3	4	4	4	3	5	2	4	4	2	1	4	5	5	2	5	3	5	5	3	5	2	5
B7-pr	4	3	5	2	2	4	4	2	4	4	1	1	4	4	4	2	4	3	4	4	4	4	4	4
B7-po	4	2	4	2	2	4	4	2	4	3	2	1	4	4	4	2	4	3	3	4	4	3	4	4
B8-pr	5	4	4	3	3	3	3	2	2	1	1	1	3	5	5	1	3	1	4	3	2	3	2	5
B8-po	4	2	3	2	2	4	4	3	4	2	1	1	3	4	5	1	5	3	3	3	2	3	1	4

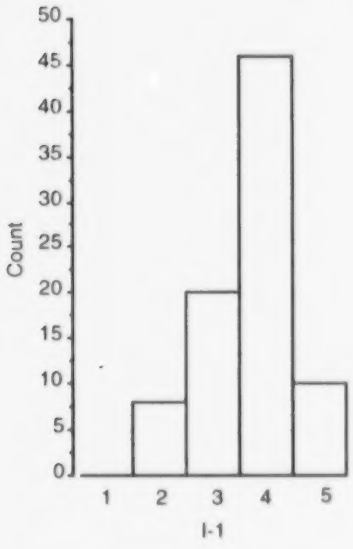
B9-pr	3	2	2	5	3	2	4	4	5	2	5	1	4	4	4	2	3	2	3	4	4	3	4	4
B9-po	3	2	2	4	4	2	4	2	4	2	2	1	4	4	4	2	3	2	4	4	4	4	2	2
B10-pr	4	2	1	1	3	5	4	2	1	3	1	1	3	4	4	1	4	3	4	2	4	3	2	2
B10-po	4	2	3	2	2	4	2	2	4	3	1	1	2	4	4	1	3	2	5	4	4	4	2	3
B11-pr	4	2	5	1	1	4	4	2	2	5	1	1	4	4	3	2	5	2	5	2	2	5	1	5
B11-po	4	2	5	2	2	5	4	2	1	5	1	1	4	3	3	1	5	2	4	3	2	5	2	5
B12-pr	3	2	4	1	2	4	4	2	3	3	1	1	5	5	4	1	4	3	3	3	4	3	2	3
B12-po	4	2	3	1	3	4	5	2	3	2	1	1	4	4	3	1	3	2	5	3	4	4	2	4
B13-pr	4	2	5	2	1	2	4	2	4	5	1	1	3	5	4	1	4	2	4	5	3	4	2	4
B13-po	3	2	4	2	1	3	3	2	2	4	1	1	2	4	4	1	4	2	4	5	3	4	2	4
B14-pr	3	1	4	2	3	3	4	2	4	2	3	1	2	4	5	1	5	2	3	4	4	4	2	4
B14-po	2	2	4	1	1	3	3	2	1	4	1	1	3	4	4	1	4	3	3	4	4	3	2	3
B15-pr	3	2	5	2	2	4	5	3	1	4	1	1	4	4	4	1	3	2	4	5	4	4	2	2
B15-po	4	3	4	2	2	4	4	2	1	4	1	1	4	4	4	1	4	1	4	4	3	4	3	4
B16-pr	4	4	3	4	2	4	4	2	1	4	1	1	4	4	4	2	4	2	4	4	2	2	4	4
B16-po	4	4	4	5	4	4	4	4	1	4	1	1	4	4	4	1	4	4	4	4	4	3	4	4
B17-pr	3	3	4	2	2	4	4	2	4	2	1	1	3	4	4	2	4	3	4	4	4	4	3	4
B17-po	4	3	4	3	3	4	4	2	3	4	1	1	4	4	4	2	4	2	4	4	3	4	2	2
B18-pr	2	3	4	2	3	4	5	3	4	5	3	1	4	4	4	4	4	3	4	5	5	5	4	4
B18-po	3	3	3	3	3	4	5	4	4	4	3	1	4	4	4	4	4	3	4	4	4	4	4	4
B19-pr	4	2	4	2	1	4	4	2	3	4	2	1	4	4	4	2	4	3	4	4	4	4	2	4
B19-po	4	3	4	5	3	4	4	2	4	2	2	1	4	4	4	2	4	3	4	4	3	4	1	4
B20-pr	4	2	5	2	4	5	4	1	3	5	1	1	1	4	4	1	5	1	3	4	3	5	2	2
B20-po	4	2	5	1	2	5	4	2	4	5	1	1	2	5	3	1	5	1	4	5	3	5	2	1
B21-pr	2	3	4	3	4	2	4	3	4	4	2	1	2	3	3	1	2	3	2	4	4	4	4	4
B21-po	4	3	5	2	3	3	5	2	3	3	2	1	4	3	4	1	3	3	4	5	3	4	5	3
B22-pr	3	5	1	4	4	3	3	1	5	3	5	1	1	3	1	1	2	2	5	5	5	5	1	5
B22-po	5	5	2	3	3	3	2	1	5	1	1	1	2	3	2	3	3	1	5	4	5	5	1	5
B23-pr	4	4	3	5	3	4	4	3	2	2	1	1	4	3	4	2	4	2	4	4	4	4	1	4
B23-po	4	3	4	4	3	4	5	3	4	3	2	1	4	4	4	2	4	3	4	4	4	4	2	4
B24-pr	5	3	4	1	2	5	5	2	3	2	1	1	2	4	5	1	5	2	2	4	4	5	2	5
B24-po	4	2	4	2	2	4	4	2	4	2	3	1	3	4	5	1	5	4	5	4	4	4	2	5
B25-pr	4	2	5	2	3	5	5	2	5	4	1	1	4	4	4	2	5	2	4	5	4	4	4	4
B25-po	4	2	4	2	2	4	5	3	4	5	2	1	4	5	4	2	5	4	4	5	3	5	4	4
B26-pr	3	2	4	1	4	4	4	1	1	4	1	1	3	3	4	1	4	3	4	4	4	4	2	5
B26-po	4	4	5	4	4	4	4	2	1	4	1	1	4	4	4	2	4	4	4	4	4	4	1	4
B27-pr	4	3	5	1	2	3	4	2	4	5	1	1	5	5	5	1	5	2	5	5	4	5	2	5
B27-po	3	2	5	4	3	1	4	2	1	5	2	1	5	5	5	1	5	2	5	4	5	5	1	5
B28-pr	3	2	5	2	1	4	5	2	2	4	1	1	4	5	5	3	5	3	4	5	5	5	4	5
B28-po	4	2	5	2	2	4	5	4	4	5	1	1	5	5	5	4	5	4	4	5	5	5	3	5
B29-pr	4	3	4	5	4	3	2	2	5	2	3	1	3	5	4	1	4	3	4	4	3	5	2	5
B29-po	4	3	4	5	5	3	3	2	4	2	2	1	4	5	4	2	4	3	4	4	4	4	3	5

Individual Subject Pretest-Posttest Differences

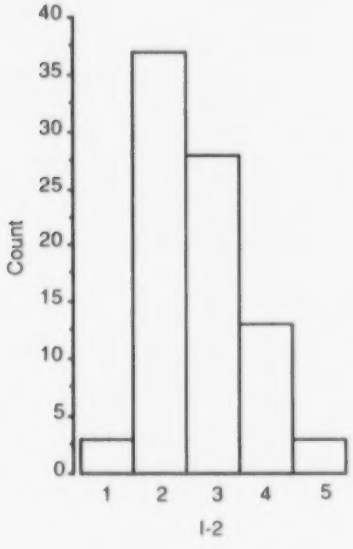
Subjects	I-1	I-2	I-3*	I-4	I-5	I-6*	II-1*	II-2	II-3	II-4*	II-5	II-6	III-1	III-2	III-3*	III-4	III-5*	III-6	IV-1	IV-2*	IV-3	IV-4*	IV-5	IV-6
A1	0	2	0	0	1	0	-1	-3	0	-2	-1	0	0	1	0	-1	0	0	-1	1	3	0	-2	-1
A2	-1	-1	0	0	-1	-2	0	0	-1	-1	1	0	0	1	1	1	0	0	0	1	0	2	-1	0
A3	0	0	1	-1	0	0	0	-2	2	0	-1	0	1	0	0	0	0	-2	0	0	0	0	0	0
A4	-1	0	0	-2	-3	0	3	-2	1	0	0	0	1	1	1	0	1	1	1	0	0	1	0	-1
A5	0	0	0	-1	0	-1	0	-1	0	0	-1	0	2	-2	0	1	1	0	0	1	0	2	-1	1
A6	0	0	0	1	1	-1	-1	-1	0	-1	0	-1	-1	1	0	0	1	0	0	-1	1	-1	1	0
A7	0	3	-1	3	4	1	0	-2	0	-2	-1	0	1	0	0	0	0	-1	0	0	2	-2	0	0
A8	-1	0	-1	0	-2	-1	0	-1	-1	0	-1	0	-2	-1	0	0	1	0	1	-1	0	0	-1	1
A9	2	-2	1	-1	1	-1	2	-2	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1
A10	1	-1	2	-1	1	0	1	0	1	0	0	0	1	0	1	0	2	1	2	0	2	0	0	2
A11	1	0	0	0	0	0	0	-2	0	0	1	0	0	0	0	0	0	0	1	1	1	1	0	1
A12	0	0	0	-1	0	0	-2	0	0	0	-2	0	-1	-1	-3	-1	0	-1	0	0	-1	-1	-1	-1
A13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	-1	-1	0	-1	-1	0	0	0
B1	0	0	-1	0	0	1	0	0	1	0	-1	-1	0	0	0	1	0	1	1	1	0	0	0	0
B2	0	-1	0	-2	-1	0	0	0	0	0	0	0	0	1	0	0	0	1	-1	0	1	0	0	-1
B3	0	-1	2	-2	0	0	0	0	0	-1	-1	0	0	-1	1	0	0	-1	0	0	-1	0	0	1
B4	1	2	0	-2	0	1	0	0	1	0	-1	-1	0	0	0	0	2	1	0	0	0	0	1	-1
B5	0	1	0	0	0	0	0	-1	1	1	1	0	-1	-1	1	1	1	0	1	1	0	1	-1	0
B6	-1	1	1	-1	0	0	-1	0	0	1	0	0	0	0	0	0	0	0	-1	0	0	0	1	0
B7	0	1	1	0	0	0	0	0	0	1	-1	0	0	0	0	0	0	0	1	0	0	1	0	0
B8	1	2	1	1	1	-1	-1	-1	-2	-1	0	0	0	1	0	0	-2	-2	1	0	0	0	1	1
B9	0	0	0	1	-1	0	0	2	1	0	3	0	0	0	0	0	0	0	-1	0	0	-1	2	2
B10	0	0	-2	-1	1	1	2	0	-3	0	0	0	1	0	0	0	1	1	-1	-2	0	-1	0	-1
B11	0	0	0	-1	-1	-1	0	0	1	0	0	0	0	1	0	1	0	0	1	-1	0	0	-1	0
B12	-1	0	1	0	-1	0	-1	0	0	1	0	0	1	1	1	0	1	1	-2	0	0	-1	0	-1
B 13	1	0	1	0	0	-1	1	0	2	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0
B14	1	-1	0	1	2	0	1	0	3	-2	2	0	-1	0	1	0	1	-1	0	0	0	1	0	1
B15	-1	-1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	-1	1	0	1	1	0	-1	-2
B16	0	0	-1	-1	-2	0	0	-2	0	0	0	0	0	0	0	1	0	-2	0	0	-2	-1	0	0
B17	-1	0	0	-1	-1	0	0	0	1	-2	0	0	-1	0	0	0	0	1	0	0	1	0	1	2
B18	-1	0	1	-1	0	0	0	-1	0	1	0	0	0	0	0	0	0	0	0	1	1	1	0	0
B19	0	-1	0	-3	-2	0	0	0	-1	2	0	0	0	0	0	0	0	0	0	0	1	0	1	0
B20	0	0	0	1	2	0	0	-1	-1	0	0	0	-1	-1	1	0	0	0	-1	-1	0	0	0	1
B21	-2	0	-1	1	1	-1	-1	1	1	1	0	0	-2	0	-1	0	-1	0	-2	-1	1	0	-1	1
B22	-2	0	-1	1	1	0	1	0	0	2	4	0	-1	0	-1	-2	-1	1	0	1	0	0	0	0
B23	0	1	-1	1	0	0	-1	0	-2	-1	-1	0	0	-1	0	0	0	-1	0	0	0	0	-1	0
B24	1	1	0	-1	0	1	1	0	-1	0	-2	0	-1	0	0	0	0	-2	-3	0	0	1	0	0
B25	0	0	1	0	1	1	0	-1	1	-1	-1	0	0	-1	0	0	0	-2	0	0	1	-1	0	0
B26	-1	-2	-1	-3	0	0	0	-1	0	0	0	0	-1	-1	0	-1	0	-1	0	0	0	0	1	1
B27	1	1	0	-3	-1	2	0	0	3	0	-1	0	0	0	0	0	0	0	0	1	-1	0	1	0
B28	-1	0	0	0	-1	0	0	-2	-2	-1	0	0	-1	0	0	-1	0	-1	0	0	0	0	1	0
B29	0	0	0	0	-1	0	-1	0	1	0	1	0	-1	0	0	-1	0	0	0	0	-1	1	-1	0

Appendix D

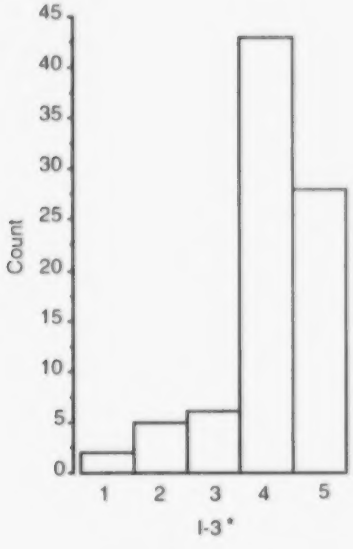
Scenario I. Frequency Distribution



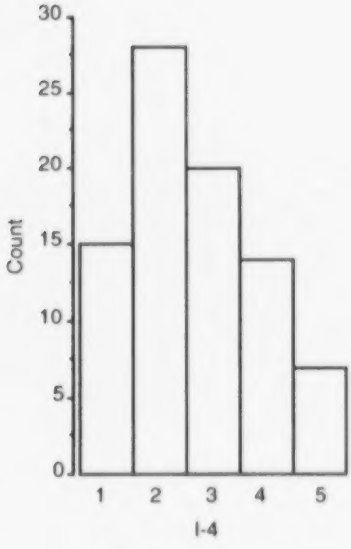
Mean: 3.69 Std. Dev.: .806



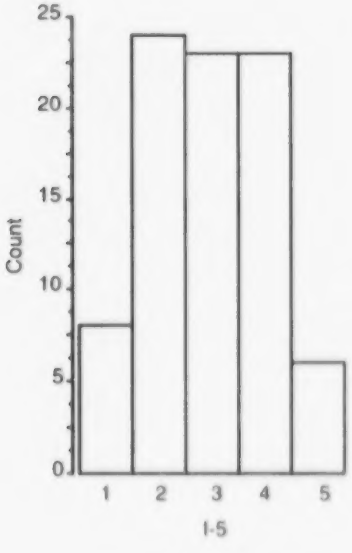
Mean: 2.714 Std. Dev.: .899



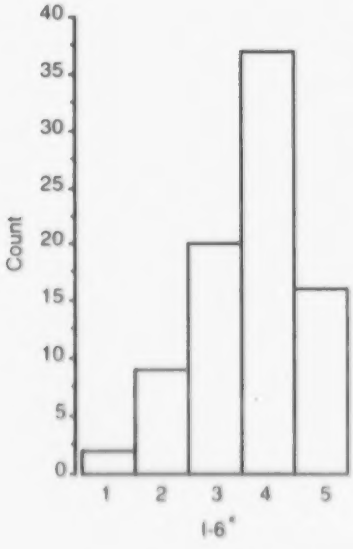
Mean: 4.071 Std. Dev.: .929



Mean: 2.643 Std. Dev.: 1.199



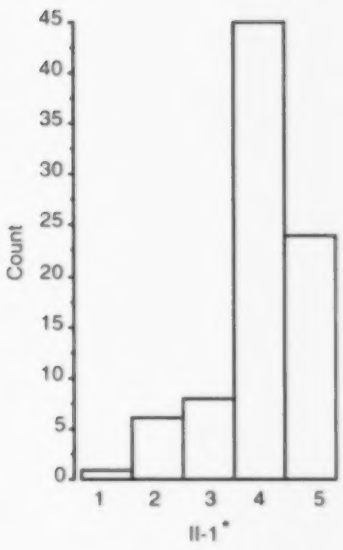
Mean: 2.94 Std. Dev.: 1.112



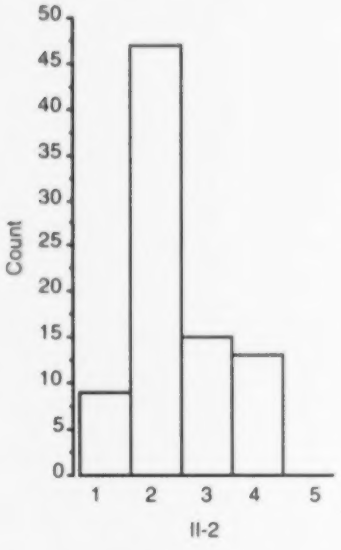
Mean: 3.667 Std. Dev.: .986

* Statements made by native speakers of English

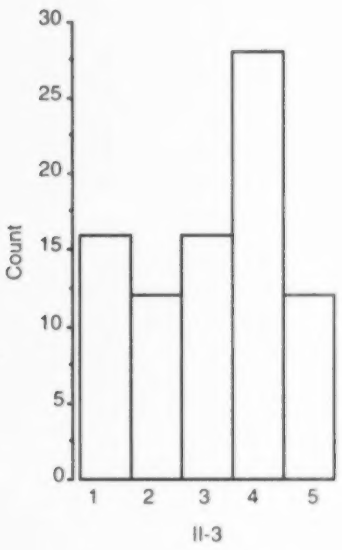
Scenario II. Frequency Distribution



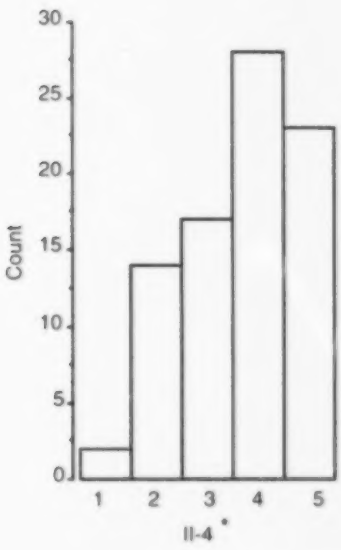
Mean: 4.012 Std. Dev.: .885



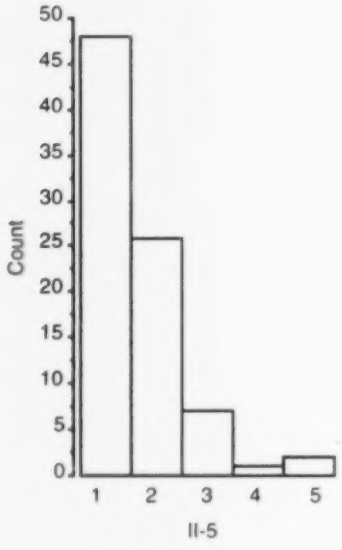
Mean: 2.381 Std. Dev.: .877



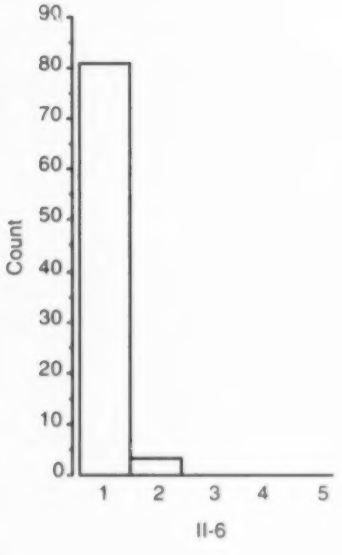
Mean: 3.095 Std. Dev.: 1.35



Mean: 3.667 Std. Dev.: 1.123



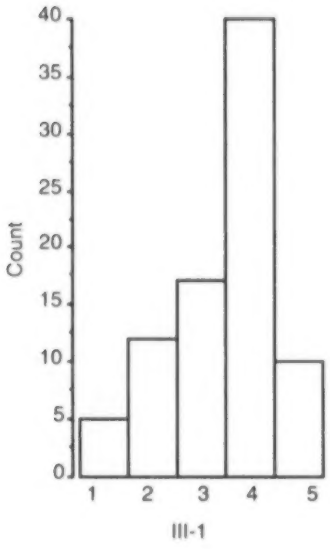
Mean: 1.607 Std. Dev.: .878



Mean: 1.036 Std. Dev.: .187

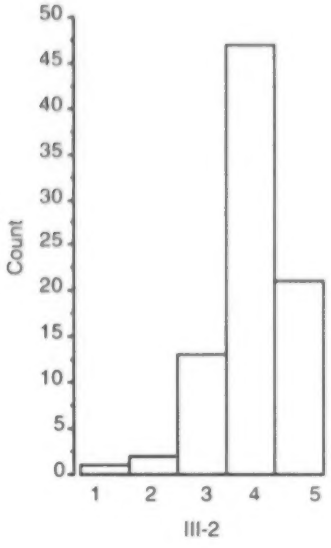
* Statements made by native speakers of English

Scenario III. Frequency Distribution



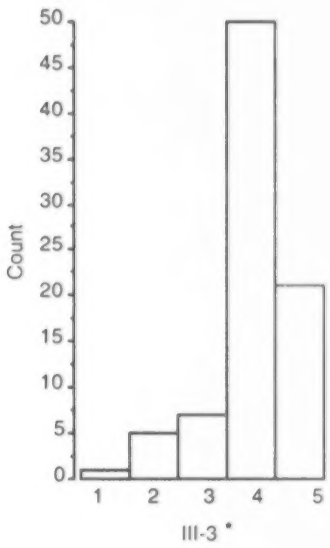
III-1

Mean:	Std. Dev.:
3.452	1.069



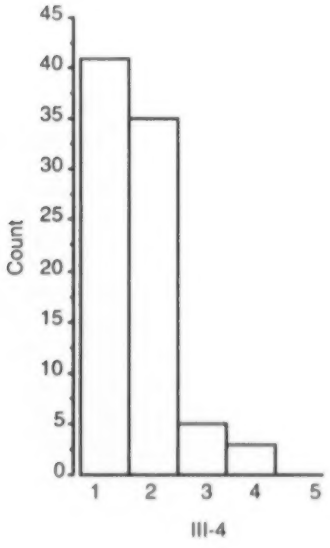
III-2

Mean:	Std. Dev.:
4.012	.784



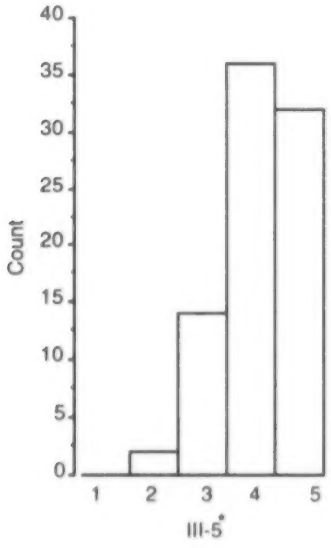
III-3 *

Mean:	Std. Dev.:
4.012	.829



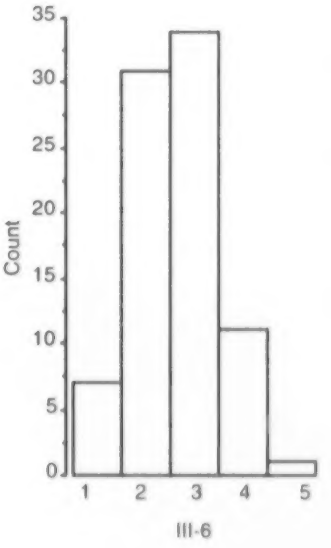
III-4

Mean:	Std. Dev.:
1.643	.755



III-5*

Mean:	Std. Dev.:
4.167	.789

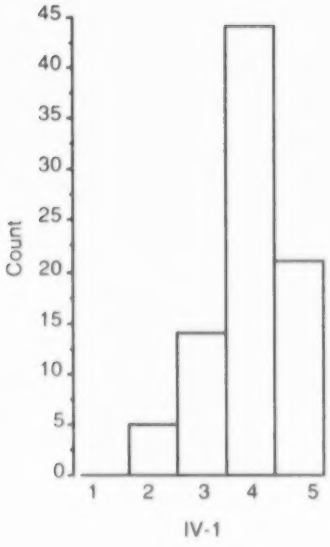


III-6

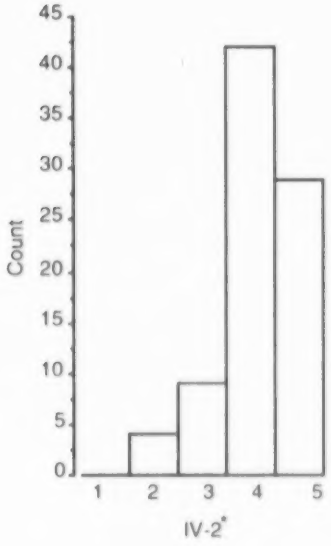
Mean:	Std. Dev.:
2.619	.863

* Statements made by native speakers of English

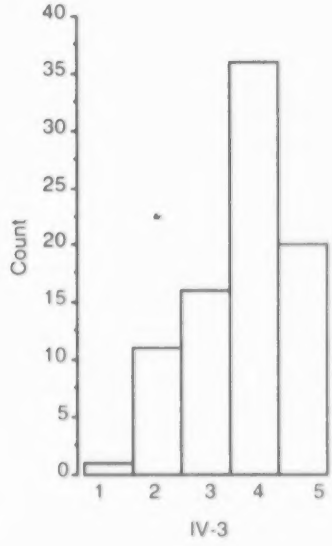
Scenario IV. Frequency Distribution



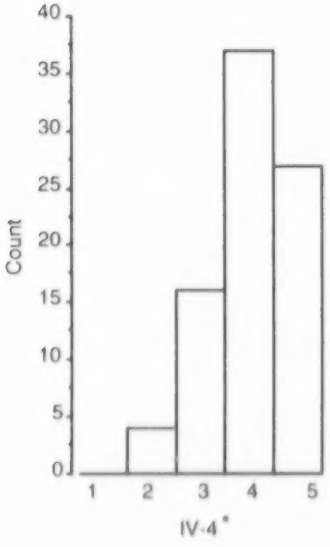
Mean:	Std. Dev.:
3.964	.813



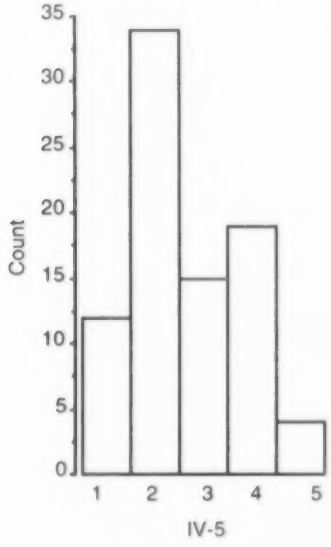
Mean:	Std. Dev.:
4.143	.794



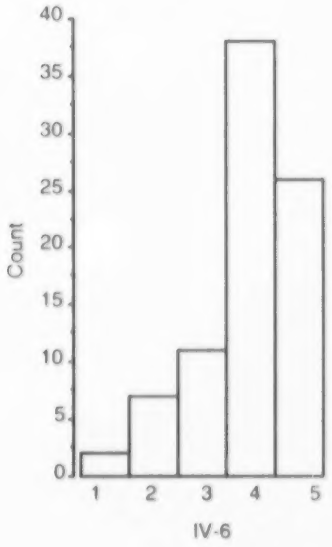
Mean:	Std. Dev.:
3.75	1.005



Mean:	Std. Dev.:
4.036	.842



Mean:	Std. Dev.:
2.631	1.128



Mean:	Std. Dev.:
3.94	.998

* Statements made by native speakers of English

Appendix E

Statement Means and Standard Deviations

Statements	Mean	Descending Rank	SD	Ascending Rank
I-1	3.69	2	0.81	1
I-2	2.71	5	0.90	2
I-3*	4.07	1	0.93	3
I-4	2.64	6	1.20	6
I-5	2.94	4	1.11	5
I-6*	3.67	3	0.99	4
II-1*	4.01	1	0.89	4
II-2	2.38	4	0.88	2
II-3	3.10	3	1.35	6
II-4*	3.67	2	1.12	5
II-5	1.61	5	0.88	3
II-6	1.04	6	0.19	1
III-1	3.45	4	1.07	6
III-2	4.01	2 / 3	0.78	2
III-3*	4.01	2 / 3	0.83	4
III-4	1.64	6	0.76	1
III-5*	4.17	1	0.79	3
III-6	2.62	5	0.86	5
IV-1	3.96	3	0.81	2
IV-2*	4.14	1	0.79	1
IV-3	3.75	5	1.01	5
IV-4*	4.04	2	0.84	3
IV-5	2.63	6	1.13	6
IV-6	3.94	4	1.00	4

** Statements made by native speakers of English*

Appendix F

Statement Rating Consistency

Statements	± 0		± 1		± 2		± 3		± 4			Total
I-1	21	50%	18	43%	3	7%		0%		0%		100%
I-2	23	55%	13	31%	5	12%	1	2%		0%		100%
I-3*	21	50%	18	43%	3	7%		0%		0%		100%
I-4	14	33%	20	48%	4	10%	4	10%		0%		100%
I-5	17	40%	18	43%	5	12%	1	2%	1	2%		100%
I-6*	26	62%	14	33%	2	5%		0%		0%		100%
II-1*	24	57%	14	33%	3	7%	1	2%		0%		100%
II-2	22	52%	11	26%	8	19%	1	2%		0%		100%
II-3	17	40%	17	40%	5	12%	3	7%		0%		100%
II-4*	22	52%	14	33%	6	14%		0%		0%		100%
II-5	21	50%	16	38%	3	7%	1	2%	1	2%		100%
II-6	39	93%	3	7%		0%		0%		0%		100%
III-1	21	50%	18	43%	3	7%		0%		0%		100%
III-2	24	57%	17	40%	1	2%		0%		0%		100%
III-3*	31	74%	10	24%		0%	1	2%		0%		100%
III-4	29	69%	12	29%	1	2%		0%		0%		100%
III-5*	27	64%	12	29%	3	7%		0%		0%		100%
III-6	19	45%	18	43%	5	12%		0%		0%		100%
IV-1	24	57%	14	33%	3	7%	1	2%		0%		100%
IV-2*	25	60%	16	38%	1	2%		0%		0%		100%
IV-3	24	57%	14	33%	3	7%	1	2%		0%		100%
IV-4*	24	57%	15	36%	3	7%		0%		0%		100%
IV-5	21	50%	19	45%	2	5%		0%		0%		100%
IV-6	21	50%	17	40%	4	10%		0%		0%		100%
Average	23.2	55.3%	14.9	35.5%	3.2	7.5%	0.6	1.5%	0.1	0.2%		100%

** -- Statements made by native speakers of English*

Appendix G

Subject Judgement Consistency

Subj.	± 0		± 1		± 2		± 3		± 4			Total
A1	11	46%	8	33%	3	13%	2	8%		0%		100%
A2	11	46%	11	46%	2	8%		0%		0%		100%
A3	17	71%	4	17%	3	13%		0%		0%		100%
A4	10	42%	10	42%	2	8%	2	8%		0%		100%
A5	12	50%	9	38%	3	13%		0%		0%		100%
A6	10	42%	14	58%		0%		0%		0%		100%
A7	12	50%	5	21%	4	17%	2	8%	1	4%		100%
A8	10	42%	12	50%	2	8%		0%		0%		100%
A9	14	58%	6	25%	4	17%		0%		0%		100%
A10	10	42%	9	38%	5	21%		0%		0%		100%
A11	16	67%	7	29%	1	4%		0%		0%		100%
A12	12	50%	9	38%	2	8%	1	4%		0%		100%
A13	19	79%	5	21%		0%		0%		0%		100%
B1	15	63%	9	38%		0%		0%		0%		100%
B2	16	67%	7	29%	1	4%		0%		0%		100%
B3	14	58%	8	33%	2	8%		0%		0%		100%
B4	13	54%	8	33%	3	13%		0%		0%		100%
B5	10	42%	14	58%		0%		0%		0%		100%
B6	16	67%	8	33%		0%		0%		0%		100%
B7	18	75%	6	25%		0%		0%		0%		100%
B8	8	33%	12	50%	4	17%		0%		0%		100%
B9	15	63%	5	21%	3	13%	1	4%		0%		100%
B10	11	46%	9	38%	3	13%	1	4%		0%		100%
B11	15	63%	9	38%		0%		0%		0%		100%
B12	11	46%	12	50%	1	4%		0%		0%		100%
B13	16	67%	7	29%	1	4%		0%		0%		100%
B14	10	42%	10	42%	3	13%	1	4%		0%		100%
B15	13	54%	10	42%	1	4%		0%		0%		100%
B16	16	67%	4	17%	4	17%		0%		0%		100%
B17	14	58%	8	33%	2	8%		0%		0%		100%
B18	16	67%	8	33%		0%		0%		0%		100%
B19	17	71%	4	17%	2	8%	1	4%		0%		100%
B20	14	58%	9	38%	1	4%		0%		0%		100%
B21	7	29%	14	58%	3	13%		0%		0%		100%
B22	11	46%	9	38%	3	13%		0%	1	4%		100%
B23	14	58%	9	38%	1	4%		0%		0%		100%
B24	13	54%	8	33%	2	8%	1	4%		0%		100%
B25	13	54%	10	42%	1	4%		0%		0%		100%
B26	13	54%	9	38%	1	4%	1	4%		0%		100%
B27	14	58%	7	29%	1	4%	2	8%		0%		100%
B28	15	63%	7	29%	2	8%		0%		0%		100%
B29	15	63%	9	38%		0%		0%		0%		100%
Average	12.6	52.6%	8.4	34.9%	2.4	9.9%	0.5	2.2%	0.1	0.3%		100%

Appendix H

Order Effect

Statements	± 0	± 1	± 2	± 3	± 4	Total of Difference††
I -1 to 6	40	27	7	3	1	54
III - 1 to 6	42	30	5	1		43
II - 1 to 6†	44	21	11	2		49
IV - 1 to 6†	38	31	8	1		50

† Orders of statements were randomized within Scenario II and IV in post-test

*†† Culculated by multiplying point of difference (i.e. 0, 1, 2, 3, or 4) by its number of occurance and then summing the total of point of difference. e.g., The total of difference for I - 1 to 6 is SUM(40*0+27*1+7*2+3*3+1*4)*